Une vision géométrique de la méthode de Ferrari pour résoudre les équations de degré 4

1 La méthode originelle de Ferrari

On connaît la méthode de Cardan (publiée dans son *Ars magna* en 1545) pour résoudre les équations de degré 3 par radicaux (voir par exemple `http://www.math.u-psud.fr/~perrin/CAPES/algebre/Cardan10.pdf`). Dans la foulée de Cardan, l'un de ses élèves, Ferrari, proposa une méthode pour résoudre l'équation de degré 4 en la ramenant à une équation de degré 3 et deux de degré 2. La méthode de Ferrari, comme celle de Cardan, repose sur une astuce de calcul qui s'éclaire (notamment l'apparition de l'équation de degré 3) en utilisant ce qu'on appelle les pinceaux de coniques.

1.1 L'astuce de Ferrari

La situation est la suivante. Soit $F(x) = 0$ une équation de degré 4, disons à coefficients réels, $F(x) = \alpha x^4 + \beta x^3 + \gamma x^2 + \delta x + \epsilon$. On suppose que cette équation admet 4 racines distinctes [1] (réelles ou complexes). On sait qu'en posant $X = x + \frac{\beta}{4\alpha}$ on élimine le terme de degré 3 de l'équation et on se ramène à une équation de la forme $P(x) = x^4 + px^2 + qx + r = 0$. Ferrari propose de comparer $P(x)$ à $Q(x)^2$ avec $Q(x) = x^2 + \lambda$:

$$P(x) = (x^2 + \lambda)^2 + x^2(p - 2\lambda) + qx + r - \lambda^2 = Q(x)^2 + T(x),$$

puis de déterminer λ de telle sorte que le deuxième terme $T(x) = x^2(p-2\lambda) + qx + r - \lambda^2$ soit l'opposé du carré d'un polynôme $S(x)$ de degré 1 en x. Si tel est le cas, on a $P(x) = Q(x)^2 - S(x)^2 = (Q(x)+S(x))\,(Q(x)-S(x))$ et l'équation $P(x) = 0$ se ramène à deux équations du second degré $Q(x) + S(x) = 0$ et $Q(x) - S(x) = 0$.

Mais on sait que pour que le polynôme de degré 2, $(2\lambda - p)x^2 - qx + \lambda^2 - r$ soit le carré d'un [2] polynôme de degré 1 il faut et il suffit qu'il ait une racine double, donc que son discriminant soit nul. Or, ici, on a

$$\Delta = q^2 - 4(2\lambda - p)(\lambda^2 - r) = 0 \iff 8\lambda^3 - 4p\lambda^2 - 8r\lambda + 4pr - q^2 = 0 := R(\lambda).$$

On voit que λ est solution d'une équation du troisième degré $R(\lambda) = 0$ où R est ce qu'on appelle une **résolvante** de P, que l'on résout par exemple par la méthode de Cardan. Une fois le λ trouvé, il reste à résoudre les deux équations

1. Si F admet des racines doubles, celles-ci annulent la dérivée de F et on est ramené à chercher les racines d'un polynôme de degré ≤ 3.

2. Si l'on veut que $S(x)$ soit réel il faut aussi que le coefficient $2\lambda - p$ soit positif.

du second degré ci-dessus : on a bien ramené la résolution de l'équation de degré 4 à celles d'une équation de degré 3 et de deux équations de degré 2.

2 La méthode avec les coniques

On reprend l'équation $F(x) = 0$ de degré 4, avec $F(x) = \alpha x^4 + \beta x^3 + \gamma x^2 + \delta x + \epsilon$ et on suppose encore que cette équation admet 4 racines distinctes.

On l'interprète comme l'équation aux x de l'intersection de deux coniques du plan affine, l'une d'équation $f(x, y) = y - x^2 = 0$ et l'autre (par exemple) $g(x, y) = \alpha y^2 + \beta xy + \gamma y + \delta x + \epsilon = 0$. En effet, en remplaçant y par x^2 dans $g(x, y) = 0$ on tombe exactement sur l'équation $F(x) = 0$. Comme les racines de F sont distinctes, ces coniques ont exactement 4 points en [3]commun, notés A, B, C, D.

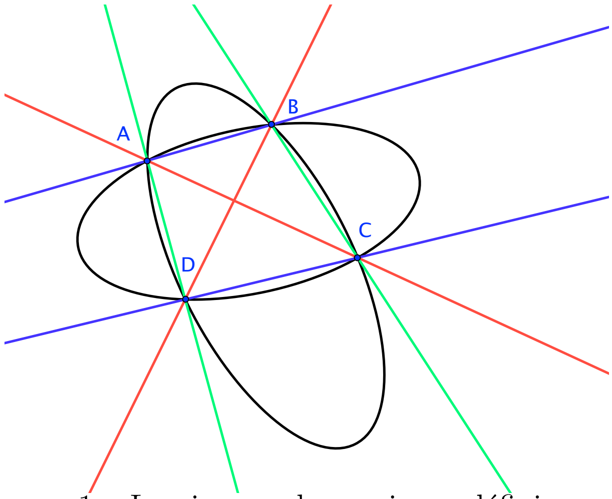


Figure 1 – Le pinceau des coniques défini par f, g

On considère alors ce qu'on appelle le **pinceau** (ou faisceau) de coniques défini par f et g. Ce sont les coniques d'équations $\lambda f + \mu g = 0$ avec $\lambda, \mu \in \mathbf{R}$ (voire $\mathbf{C}$) ou même $\lambda f + g$ si l'on exclut la conique définie par f. Ces coniques passent par les quatre points A, B, C, D (on montre que ce sont exactement les coniques passant par ces points mais nous n'en aurons pas besoin). Parmi les coniques passant par ces quatre points, il y en a trois qui sont particulières, ce sont les coniques dégénérées en deux droites, ici $(AB) \cup (CD)$, $(AC) \cup (BD)$ et $(AD) \cup (BC)$. En effet, une droite a une équation qui est une forme affine $l(x) = ux + vy + w = 0$ et la réunion de deux droites a pour équation le

3. Le cas de la figure est celui où les quatre racines sont réelles.

produit de deux telles formes, c'est donc bien une conique. Ces coniques dégénérées vont avoir des équations de la forme $\lambda f + g = 0$ correspondant à trois valeurs particulières de λ.

La méthode de Ferrari (revisitée) va consister à déterminer ces valeurs de λ (on verra qu'elles sont racines d'une équation de degré 3). Une fois en possession de l'équation d'une de ces coniques dégénérées, disons $(AB) \cup (CD)$, on trouvera les équations des deux droites (AB) et (CD) et il ne restera plus, pour trouver les abscisses de A, B, C, D qu'à couper l'une des coniques par ces deux droites, donc à résoudre deux équations de degré 2, pour trouver les racines de F.

3 Le critère de décomposition

3.1 L'énoncé

Le point crucial de la méthode est de disposer d'un critère permettant d'affirmer qu'une conique est décomposée en deux droites :

3.1 Théorème. *On considère un polynôme de degré* 2 *en* x, y *:*

$$P(x,y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f.$$

Alors, P *est produit de deux polynômes du premier degré si et seulement si on a :*

$$acf + 2bde - ae^2 - fb^2 - cd^2 = 0.$$

3.2 La preuve matricielle

Lorsqu'on connaît un peu la théorie des coniques (voir par exemple `http://www.math.u-psud.fr/~perrin/Livregeometrie/DPPartie3.pdf`, on sait que cette condition n'est autre que l'annulation du déterminant $\begin{vmatrix} a & b & d \\ b & c & e \\ d & e & f \end{vmatrix}$ qui est le discriminant de la forme quadratique obtenue à partir de P par homogénéisation :

$$ax^2 + 2bxy + cy^2 + 2dxt + 2eyt + ft^2.$$

3.3 Une preuve directe pour le sens facile

Il s'agit de montrer que si l'on a la relation $acf + 2bde - ae^2 - fb^2 - cd^2 = 0$, le polynôme P est produit de deux formes linéaires. Comme P est de degré

2, les coefficients a, b, c ne sont pas tous nuls. Le cas $a = c = 0$ est facile et laissé au lecteur (on écrit $P = (2bx + 2e)(y + d/b)$). Si a ou c est non nul, quitte à échanger les rôles de x et y, on peut supposer $a \neq 0$ et, en divisant par a, on se ramène au cas $a = 1$. On utilise alors la même ruse que Ferrari en écrivant :

$$P(x) = (x + by + d)^2 - [(b^2 - c)y^2 + 2(bd - e)y + d^2 - f]$$

et P sera factorisé si le second terme est un carré, ce qui sera le cas si son discriminant est nul. Or, on a :

$$\Delta' = (bd - e)^2 - (b^2 - c)(d^2 - f) = -cf - 2bde + e^2 + fb^2 + cd^2$$

et cette quantité est nulle par hypothèse.

3.4 La preuve "géométrique" de 3.1

Supposons d'abord que P est produit de deux polynômes Q, R de degré 1 qui correspondent donc à des équations de droites. On a alors le lemme suivant (dans lequel on écrit simplement que M est un point singulier de P) :

3.2 Lemme. *Le point d'intersection M des droites d'équations $Q = R = 0$ est donné par les équations : $ax+by+d = 0$ et $bx+cy+e = 0$. Ses coordonnées sont donc :*

$$x_0 = \frac{cd - be}{b^2 - ac}, \quad y_0 = \frac{ae - bd}{b^2 - ac}.$$

Démonstration. On dérive P par rapport à x, y, on a :

$$\frac{\partial P}{\partial x}(x, y) = 2ax + 2by + 2d = \frac{\partial Q}{\partial x}(x, y)R(x, y) + \frac{\partial R}{\partial x}(x, y)Q(x, y)$$

$$\frac{\partial P}{\partial y}(x, y) = 2bx + 2cy + 2e = \frac{\partial Q}{\partial y}(x, y)R(x, y) + \frac{\partial R}{\partial y}(x, y)Q(x, y).$$

On voit que le point d'intersection de Q et R annule bien les polynômes annoncés.

Comme on a $P = QR$, le point M est sur P et on conclut grâce au lemme suivant :

3.3 Lemme. *Le point $M = (x_0, y_0)$ vérifie $P(x_0, y_0) = 0$ si et seulement si on a $dx_0 + ey_0 + f = 0$, c'est-à-dire $acf + 2bde - ae^2 - fb^2 - cd^2 = 0$.*

Démonstration. Cela vient de la formule évidente suivante :

$$P(x,y) = x(ax+by+d) + y(bx+cy+e) + dx + ey + f.$$

3.4 *Remarque.* Cette formule est une conséquence de la formule d'Euler, pour un polynôme homogène de degré n on a :

$$nP(x,y,t) = x\frac{\partial P}{\partial x} + y\frac{\partial P}{\partial y} + t\frac{\partial P}{\partial t}.$$

3.5 *Remarque.* Le lecteur scrupuleux mais pénible qui ferait remarquer que, dans la preuve précédente, on a supposé que les droites d'équations Q, R se coupent (ou encore que $ac - b^2$ est supposé non nul), écoperait de deux heures de colle et de la punition consistant à prouver que, si ces droites sont parallèles, on a $P(x,y) = (\alpha x + \beta y + \gamma)(\alpha x + \beta y + \delta)$ et que la condition est encore vérifiée.

Il reste à montrer la réciproque. Appelons encore $M = (x_0, y_0)$ le point défini en 3.2. Par définition de M, on a $\frac{\partial P}{\partial x}(x_0,y_0) = \frac{\partial P}{\partial y}(x_0,y_0) = 0$ et la condition du théorème signifie qu'on a aussi $P(x_0,y_0) = 0$. En effectuant le changement de variables $X = x - x_0$, $Y = y - y_0$, on peut écrire P sous la forme suivante[4] :

$$P(x,y) = \alpha(x-x_0)^2 + 2\beta(x-x_0)(y-y_0) + \gamma(y-y_0)^2 + 2\delta(x-x_0) + \epsilon(y-y_0) + \eta.$$

On a alors le lemme suivant, qui résulte d'un calcul sans difficulté :

3.6 Lemme. *Avec les notations précédentes, la condition $P(x_0,y_0) = 0$ (resp. $\frac{\partial P}{\partial x}(x_0,y_0) = 0$, resp. $\frac{\partial P}{\partial y}(x_0,y_0) = 0$) équivaut à $\eta = 0$ (resp. $\delta = 0$, resp. $\epsilon = 0$).*

La réciproque découle de ce lemme. Les hypothèses impliquent qu'on a :

$$P(x,y) = \alpha(x-x_0)^2 + 2\beta(x-x_0)(y-y_0) + \gamma(y-y_0)^2$$

et ce polynôme est produit de deux polynômes du premier degré. En effet, il est de la forme $G(X,Y) := \alpha X^2 + 2\beta XY + \gamma Y^2$ et il s'agit de voir qu'il est produit de deux polynômes de degré 1 en X, Y. Si $\alpha = 0$ on a $G(X,Y) = Y(2\beta X + \gamma Y)$. Si α est non nul, on appelle u, v les racines (dans $\mathbf{C}$) de $\alpha t^2 + \beta t + \gamma = 0$ et on a alors $G(X,Y) = \alpha(X - uY)(X - vY)$ comme on le voit en posant $t = X/Y$.

3.7 *Remarque.* Le lecteur pénible qui poserait la même question qu'à la remarque précédente serait cette fois renvoyé sans ménagement à l'étude de la géométrie projective, dans laquelle ce genre de cas particuliers disparait.

4. En fait, il s'agit simplement de la formule de Taylor et cela explique le lemme suivant.

4 La méthode de Ferrari revisitée

4.1 Le cas général

On applique la méthode indiquée dans la section 2 à l'équation $F(x) = \alpha x^4 + \beta x^3 + \gamma x^2 + \delta x + \epsilon = 0$ vue comme l'équation aux x de l'intersection des coniques $f(x, y) = y - x^2 = 0$ et $g(x, y) = \alpha y^2 + \beta xy + \gamma y + \delta x + \epsilon = 0$ et on note A, B, C, D les points d'intersection de ces coniques. On considère la conique d'équation $P := \lambda f + g = 0$, soit encore :

$$-\lambda x^2 + \beta xy + \alpha y^2 + \delta x + (\lambda + \gamma)y + \epsilon = 0.$$

Elle passe par les quatre points A, B, C, D. La condition pour que cette conique soit dégénérée est donnée par le théorème 3.1 appliqué à P. On voit qu'elle donne une équation de degré 3 en λ (à cause du terme ae^2) appelée encore **résolvante**.

Cette équation admet trois racines complexes que l'on calcule par la méthode de Cardan. On choisit l'une de ces racines, disons λ, et on considère la conique d'équation $\lambda f + g$, qui est décomposée en deux droites. On détermine les équations de ces droites (soit en utilisant la méthode du paragraphe précédent, soit par un calcul direct à la main en faisant apparaître le début d'un carré [5]). Il ne reste plus qu'à chercher les intersections de la conique d'équation $f = 0$ avec chacune de ces droites (donc à résoudre deux équations de degré 2) et on a gagné.

4.2 L'équation sans terme en x^3

Si l'équation initiale est de la forme $x^4 + px^2 + qx + r = 0$, la résolvante se calcule aisément et elle vaut :

$$R^*(\lambda) = \lambda^3 + 2p\lambda^2 + \lambda(p^2 - 4r) - q^2.$$

Ce n'est pas exactement celle donnée par Ferrari, qui valait $R(\lambda) = 8\lambda^3 - 4p\lambda^2 - 8r\lambda + 4pr - q^2 = 0$ mais elles sont liées par la formule $R(\lambda) = R^*(2\lambda - p)$.

5 Un exemple

L'exemple ci-dessous est truqué pour que les calculs soient faciles. Il s'agit de résoudre l'équation : $x^4 - \frac{3}{2}x^2 + 2x - \frac{3}{16} = 0$. On a donc $\alpha = 1$, $\beta =$

5. C'est exactement, après homogénéisation, la méthode de Gauss de décomposition en carrés des formes quadratiques.

0, $\gamma = -\frac{3}{2}$, $\delta = 2$ et $\epsilon = -\frac{3}{16}$. L'étude de la fonction montre que cette équation admet deux racines réelles et deux racines complexes conjuguées. On la considère comme l'équation aux x d'intersection des coniques d'équations $y = x^2$ et $y^2 - \dfrac{3}{2}y + 2x - \dfrac{3}{16} = 0$. En vertu de 3.1, les coniques dégénérées du pinceau sont données par l'équation en λ : $\lambda^3 - 3\lambda^2 + 3\lambda - 4 = 0$. Cette équation s'écrit encore $(\lambda - 1)^3 = 3$ (c'est là qu'intervient le trucage !) et donne donc $\lambda = 1 + \sqrt[3]{3}$. On choisit par exemple la racine cubique réelle et on décompose le polynôme associé :

$$y^2 - \frac{3}{2}y + 2x - \frac{3}{16} + \lambda(y - x^2) = \big(y + \frac{\lambda}{2} - \frac{3}{4}\big)^2 - \lambda(x - \frac{1}{\lambda})^2$$

(on n'oubliera pas la relation qui définit λ et peut s'écrire $\big(\dfrac{\lambda}{2} - \dfrac{3}{4}\big)^2 = \dfrac{1}{\lambda} - \dfrac{3}{16}$). On obtient ainsi deux droites d'équations :

$$y - \sqrt{\lambda}x + \frac{\lambda}{2} - \frac{3}{4} + \frac{1}{\sqrt{\lambda}} = 0 \quad \text{et} \quad y + \sqrt{\lambda}x + \frac{\lambda}{2} - \frac{3}{4} - \frac{1}{\sqrt{\lambda}} = 0.$$

que l'on coupe par $y = x^2$. Cela fournit deux équations du second degré en x :

$$x^2 - \sqrt{\lambda}\, x + \frac{\lambda}{2} - \frac{3}{4} + \frac{1}{\sqrt{\lambda}} = 0 \quad \text{et} \quad x^2 + \sqrt{\lambda}\, x + \frac{\lambda}{2} - \frac{3}{4} - \frac{1}{\sqrt{\lambda}} = 0.$$

La première admet deux racines imaginaires conjuguées, l'autre deux racines réelles :

$$x = \frac{-\sqrt{\lambda} \pm \sqrt{-\lambda + 3 + 4/\sqrt{\lambda}}}{2},$$

avec $\lambda = 1 + \sqrt[3]{3}$ (valeurs approchées : $0,1014101$ et $-1,6641799$).

6 La méthode de Lagrange

Indiquons pour terminer une autre méthode permettant de comprendre la résolution des équations de degré 4. Cette méthode a été introduite indépendamment en 1770 par Lagrange et Van der Monde. Elle repose sur l'usage des fonctions symétriques des racines et son intérêt majeur est d'annoncer la théorie de Galois. Reprenons un polynôme de la forme $P(x) = x^4 + px^2 + qx + r$ qui admet quatre racines x_1, x_2, x_3, x_4. En écrivant $P(x) = \displaystyle\prod_{i=1}^{4}(x - x_i)$, on peut identifier les coefficients de l'équation comme les fonctions symétriques des racines : $x_1 + x_2 + x_3 + x_4 = 0$, $\sum_{i<j} x_i x_j = p$, $\sum_{i<j<k} x_i x_j x_k = -q$ et $x_1 x_2 x_3 x_4 = r$. L'idée, pour trouver les x_i est de chercher une quantité λ, exprimée en fonction des x_i qui, lorsqu'on permute les x_i, ait moins de quatre transformés.

Un essai consiste à prendre $x_1 + x_2$. Hélas, il a au contraire plus de transformés $= x_1 + x_2$, $x_1 + x_3$, $x_1 + x_4$, $x_2 + x_3$, $x_2 + x_4$ et $x_3 + x_4$. Mais cet essai infructueux nous indique le bon candidat : $\lambda_1 = (x_1 + x_2)(x_3 + x_4)$ dont les seuls transformés sont $\lambda_2 = (x_1 + x_3)(x_2 + x_4)$ et $\lambda_3 = (x_1 + x_4)(x_2 + x_3)$.

Ces nombres vérifient évidemment l'équation $(\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) = 0$ et on calcule facilement les coefficients de cette équation. Par exemple, le coefficient de λ^2 qui est $-(\lambda_1 + \lambda_2 + \lambda_3)$ n'est autre que $-\sum_{i<j} \lambda_i \lambda_j = -p$. On montre ainsi que les λ_i sont racines de $R^\sharp(\lambda) = \lambda^3 - 2p\lambda^2 + (p^2 - 4r)\lambda + q^2 = R^*(-\lambda)$ et on résout cette équation.

Il reste à trouver les x_i à partir des λ_i. C'est facile car on a à la fois $(x_1 + x_2) + (x_3 + x_4) = 0$ et $\lambda_1 = (x_1 + x_2)(x_3 + x_4)$, de sorte qu'on connaît $x_1 + x_2$ et $x_3 + x_4$ et de même pour les autres. On en déduit aussitôt les x_i.

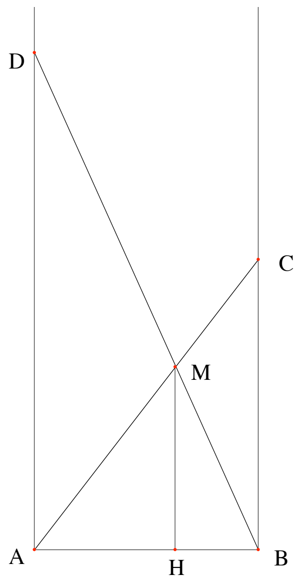
7 Une application : le problème des échelles

Deux échelles AC et BD de 2 m et 3 m sont posées dans un couloir comme sur la figure ci-contre. Le point d'intersection M est situé à 1 m du sol. On cherche à déterminer la largeur a du couloir.

1) Montrer que a est racine de l'équation $f(x) = 0$ avec

$$f(x) = \frac{1}{\sqrt{4 - x^2}} + \frac{1}{\sqrt{9 - x^2}} - 1.$$

(On calculera $\dfrac{MH}{AD}$ et $\dfrac{MH}{BC}$).



2) On pose $X = x^2$. Montrer que X est racine de l'équation de degré 4 :

$$X^4 - 22X^3 + 163X^2 - 454X + 385 = 0$$

et résoudre cette équation, d'abord par la méthode de Ferrari, puis par une méthode approchée de type Newton.

Merci à Pascal Gamblin pour ses remarques nombreuses et pertinentes.